



IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of

Appleby, et al.

Serial No.: 09/916,513

Group Art Unit: 2112

Filed: July 30, 2001

Examiner: Huynh, K.

For: METHOD, SYSTEM, AND PROGRAM PRODUCTS FOR DISTRIBUTED  
CONTENT THROTTLING IN A COMPUTING ENVIRONMENT

Honorable Commissioner of Patents  
Alexandria, Virginia 22313-1450

**DECLARATION UNDER 37 C. F. R. §1.131**

Sir:

We, Karen Appleby, Liana Liyow Fong, German Sergio Goldsmidt, Srirama  
Mandiyam Krishnakumar, and Donald Philip Pazel, do hereby state that:

- 1) We are the inventors of the above-identified application.
- 2) The method, system, and program products for distributed content throttling in a computing environment was known to us earlier than August 17, 2000, as shown in the enclosed memorandum (Exhibit 1). The effective date of the memorandum (Exhibit 1) is earlier than August 17, 2000, and the date thereof has been redacted.
- 3) The contents of the enclosed memorandum (Exhibit 1) have been incorporated into the specification of the present invention, upon which claims 1-21 and 24-33 are based. For example, the memorandum paragraph (e.g., A) describing the problem being addressed and beginning with the words "This invention relates to managing load on specific components..." can be found in the specification (e.g., on page 1, line 14 through page 2, line 13).

**BEST AVAILABLE COPY**

The memorandum paragraph (e.g., B) beginning with the words “The servers are physically inter-connected ...” can be found in the specification (e.g., on page 2, line 14 through page 3, line 4).

The memorandum paragraph (e.g., C) beginning with the words “One problem with IP layer ...” can be found in the specification (e.g., on page 3, lines 5-19).

The memorandum paragraph (e.g., D) beginning with the words “A web server farm consists ...” can be found in the specification (e.g., on page 4, lines 1-13).

4) Regarding the present invention itself, the memorandum paragraph (e.g., E) beginning with the words “This invention relates to allocating...” can be found in the specification (e.g., on page 1, lines 6-12).

The memorandum paragraph (e.g., F) beginning with the words “The present invention relates ...” can be found in the specification (e.g., on page 8, line 3 through page 9, line 8).

The memorandum paragraph (e.g., G) beginning with the words “Fig. 3 depicts the organization ...” can be found in the specification (e.g., on page 9, line 9 through page 11, line 20).

The memorandum paragraph (e.g., H) beginning with the words “Fig. 4 is a flow chart ...” can be found in the specification (e.g., on page 12, line 1 through page 13, line 13).

The memorandum paragraph (e.g., I) beginning with the words “Fig. 5 is a flow chart ...” can be found in the specification (e.g., on page 13, line 14 through page 14, line 4).

The memorandum paragraph (e.g., J) beginning with the words “Fig. 6 is a flow chart ...” can be found in the specification (e.g., on page 14, line 5 through page 15, line 14).

The memorandum paragraph (e.g., K) beginning with the words “The embodiment described above ...” can be found in the specification (e.g., on page 15, line 15 through page 16, line 2).

The first proposed claim (e.g., L) beginning with the words “A method of controlling ...” can be found in claim 1 on page 19.

5) The above clearly evidences a completion and actual reduction to practice of the invention in this country before the filing date (e.g., August 17, 2000) of Provisional Application No. 60/225,888, upon which U.S. Patent Application Publication No. US 2002/0091825 A1 is based.

6) In the alternative, we declare that the claimed invention was conceived prior to August 17, 2000 (e.g., as shown by the attached Memorandum (Exhibit 1) having a date (now redacted) prior to August 17, 2000) and, coupled with due diligence from a date before August 17, 2000, that the invention was constructively reduced to practice on July 30, 2001.

That is, a prior art search to be done internally by assignee, International Business Machines Corporation, was requested (Exhibit 2) by Douglas W. Cameron, Esq., of International Business Machines Corporation, on June 13, 2000, and the prior art search result (Exhibit 3) was reported to Mr. Cameron on July 6, 2000.

On July 12, 2000, Mr. Cameron requested comments from the inventors on the results of the internal prior art search (Exhibit 4). Internal email dialog (Exhibits 5-12) concerning the significance of the invention and its relevance with the prior art search results and clarification of the invention occurred between the inventors and Mr.

Declaration Under Rule 37 CFR §1.131  
S/N 09/916,513  
Attorney Docket YOR920010330US1

Cameron on July 31, 2000, August 1, 2000, February 2, 2001, February 13, 2001, February 22, 2001, March 15, 2001, and March 26, 2001.

Disclosure materials were forwarded (Exhibit 13) by Douglas W. Cameron, Esq. of International Business Machines Corporation to Attorney Sean M. McGinn on April 4, 2001.

An estimate for preparation of a patent application (Exhibit 14) was forwarded from Attorney Sean M. McGinn to Mr. Cameron on April 10, 2001.

Preparation of the Application was begun on May 23, 2001, by Attorney Frederick E. Cooperrider, including a telephone conference with inventor Srirama M. Krishnakumar on May 24, 2001, following which teleconference a first draft was sent to inventor Krishnakumar via email (Exhibit 15). Comments from inventor Krishnakumar were received by email on May 25, 2001 (Exhibit 16). On June 15, 2001, a second draft of the Application was sent to inventor Krishnakumar via email on June 16, 2001 (Exhibit 17). The final draft of the Application completed review at McGinn&Gibb and was sent (Exhibit 18) to the inventors via email on July 12, 2001, with formal papers.

The Application and the executed formal papers were filed on July 30, 2001.

7) The facts above in 5) clearly show a completion of the invention in the U.S. before August 17, 2000.

Alternatively, the facts in 6) above show a conception of the invention, prior to August 17, 2000, and due diligence from just before August 17, 2000, to the filing date (i.e., the constructive reduction to practice) of the Application on July 30, 2001.

We hereby declare that all statements made here are of own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the

Declaration Under Rule 37 CFR §1.131

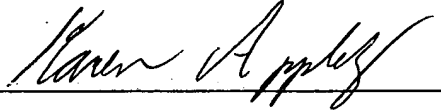
S/N 09/916,513

Attorney Docket YOR920010330US1

like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the Application or any patent issued thereon .

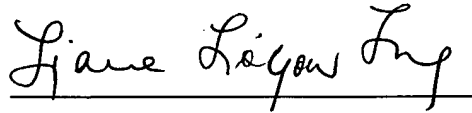
Further declarants sayeth not.

Date: Oct 8, 2004



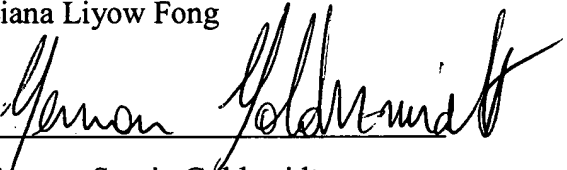
Karen Appleby

Date: Oct 7, 2004



Liana Liyow Fong

Date: Oct 7, 2004



German Sergio Goldsmidt

Date:

Srirama Mandyam Krishnakumar

Date: Oct. 7, 2004



Donald Philip Pazel

Declaration Under Rule 37 CFR §1.131

S/N 09/916,513

Attorney Docket YOR920010330US1

like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the Application or any patent issued thereon .

Further declarants sayeth not.

Date:

---

Karen Appleby

Date:

---

Liana Liyow Fong

Date:

---

German Sergio Goldsmidt

Date: 10/04/2004

*Srirama M K*  
Srirama Mandyam Krishnakumar

Date:

---

Donald Philip Pazel

## A method for distributed content based request throttling in a computing environment

Karen Appleby  
Liana Fong  
German Goldszmidt  
Srirama Krishnakumar  
Donald Pazel  
IBM Corporation

EXHIBIT 1

### Field of Invention

(E) → This invention relates to allocating internet load for web components (application components such as Enterprise Java Beans, databases, servlets etc.) in a server farm by controlling the admittance of requests based on the content of the request. The method's focus is on the utilization of the distributed processing capability inherent in the web serving environment to limit the impact on overall system performance.

### Background Description

(A) → This invention relates to managing load on specific components in a web server farm. The main function of a web server farm is to serve HTML (Hyper Text Markup Language) web pages to clients using web browsers such as Netscape or Internet Explorer. The HTML pages are generated at a component known as the web server. The web server generates the pages by processing the HTTP (Hyper Text Transfer Protocol) requests made by the web browser clients. The HTTP request consists of the web page identifier in the form of a URL (uniform request locator) and several other parameters such as cookies that could be used by the web server. The web server processes the requests by either fetching the web pages from a local repository that could be on a hard drive or by passing the request to an external component that may generate the web page. Examples of these external components are application servers, CGI (Common Gateway Interface) scripts etc. These components may be located on a different hardware platform from the web server. An example of this type of request would be a request to display the cost of items in a shopping cart. The external module in this case may interact with a backend database server to obtain the cost of the items. It then creates a HTML web page customized for this request and sends the page back to the web server. A web server farm consists of a heterogeneous mix of components such as web servers, application servers, database servers etc., that perform different functions. These components may run on different types of hardware platforms (referred to as just servers) with different capabilities (e.g amount of memory, processor speed and type etc.).

The servers are physically inter-connected by networks which are in turn connected to the Internet. Fig 1 shows an example of a web server farm. Here the servers are logically connected in a tiered manner such that the web servers (030) form the outer most tier are

(B)

followed by the application servers ~~(040)~~ and database servers ~~(050)~~. The web browsers make a TCP connection request to the web server via the Internet (010) and send the HTTP request on this connection. Some networks could have IP L4LB (layer 4 load balancing) components ~~(020)~~ which distributes the incoming TCP connections from the web browsers to the web servers that are on the network [1]. When the web browser makes a TCP connection request the L4LB redirects the request to a web server that is capable of handling the request. The web server then processes this request as explained in the paragraph above.

One problem with IP layer 4 based load balancing is that it does not look at the payload of the data and there is no differentiation among requests e.g. no differentiation between a request for a static web page that is stored in a repository such as a disk and a request to create a dynamic page by obtaining information from a backend database. This leads to the following problems when connection requests are subjected to admission control at the dispatcher level [2]. First of all requests for static pages that are relatively small and impose very little load on the system could get dropped. Second, this method could cause web site outages as the requests to overloaded components do not get sufficiently throttled (by throttling we mean a method of dropping requests when the load on the processing component increases) thereby causing them to crash.

A web server farm consists of a finite number of computing resources that are shared for processing requests from web clients. Different clients need different functions such as obtaining a static web page, obtaining a list of items in a catalog that is typically handled by a servlet or an EJB etc. These functions impose different loads on the computing resources and the functions have varying levels of importance for the business that is hosting the web site. Current state of the art treats all the requests to be of equal priority and processes the requests in the order in which they arrive. If there is an overload on the overall system requests are dropped by the L4LB. Businesses that are conducted over the internet suffer from this approach as there is no priority based dropping of the request. E.g. a customer who is requesting a web page to submit his credit number for processing is clearly more important than a customer who is requesting to view the picture of an item in the catalog. Thus, in this case the credit card processing web page has a higher priority when compared to the page that displays a catalog item.

### ***Detailed Description of A Preferred Embodiment Of The Invention***

The present invention relates to managing load on specific components in a web server farm. A web server farm consists of a heterogeneous mix of components with different capabilities and perform different functions such as a web serving, database serving, application serving etc. They are typically structured in a tiered manner and requests flow from one tier to the other. Fig 7 shows a typical web server farm. Requests for performing tasks typically arrive into the farm from the internet ~~101~~. The Internet forwards the request into the router ~~102~~, which is typically located at the entry to the farm or at a location in the Internet Service provider's point of presence. The router uses the IP address contained in the incoming packet to route the request to one of the L4LB's.

Exh. 1  
(page 2 of 6)



~~or 104~~. The L4LB provides a virtualization of the IP addresses of all the servers that are connected to it ~~(the web servers are numbered 105 through 112)~~. The L4LB distributes the TCP connection requests using weights assigned to the web servers that are connected to it. This distribution is typically done during the arrival of a TCP connection and not during each individual HTTP 1.1 request that is made on the connection. ~~113 and 115 are servlets that are used to process certain requests.~~ Servlets are written in Java and run in a Java Virtual Machine container such as Websphere application server, JServ or TomCat. 114 is a CGI bin application which also processes certain types of requests. 116 is an application server which contains components such as Enterprise Java Beans used to perform business logic functions. 117 is a database which contains data that is critical to the processing of requests served by the other components such as the servlet or application server.

Fig 3 depicts the organization of components in a web serving environment to support content based throttling. 201 is the web server request processor at which the requests arrive after they have been dispatched by the L4LB ~~(103 or 104)~~. In the preferred embodiment the web server request processor, content throttler (202) and the content handler are all implemented as part of the web server. Web servers typically provide a feature where one can plugin software components that are typically not part of the web server. These components can be interposed in the request processing phases of the web server to provide additional functionality or modify the functionality provided by the base web server. Such software modules are referred to as "plugins". The content throttler is implemented as a plugin in this embodiment. While the content handler is typically a part of the core web server, it can be replaced by a plugin that can process specific types of requests (an example of such a function would be to replace a file handler for a streaming video image with a custom page handler when the bandwidth utilization is high). The request processing in a web server typically consists of inspecting the URL of a request and determining the content handler that can process the request based on the type of the request. A request type is identified by the configuration information that is present in the web server configuration. Requests can be broadly classified into static and dynamic requests. Static requests are requests for information that is contained in a static repository such as HTML file or a GIF image file. Dynamic requests are for information that has to be constructed as a response to the request. Dynamic requests cannot be stored in a repository and presented to the user when the request has been made. Typically dynamic requests are for information such as items in the shopping cart of a client, information about the stock portfolio of the client etc. The web server configuration contains information that maps the type of request with the content handler that can process the request. The content handler processes the request by performing the necessary tasks to deliver the content requested. If the request is for a static HTML file or a GIF image, it obtains the file by making an operating system dependent call to get the file and delivers the file on the connection that has been established between the web server and the client's web browser. If the request is for obtaining the items in the shopping cart of a client, then the content handler may initiate a query to a backend database to fetch the items in the catalog, construct a HTML web page with the result of the query after formatting the web page suitably and deliver the page to the web browser.

client on the connection that has been established between the browser and the server.

Fig 4 is a flow chart for that depicts the process of content based throttling in a web serving environment. At step 301 requests arrive at the web server request processor (201) after a connection has been established. At step 302 the web server determines if the request should be subjected to content based throttling. This process is explained in the following sections and in figure 5 in detail. If it is determined that the request does not need to be throttled, then the request is handed over to the content handler. At step 304 the content handler (203) produces the content, e.g. a HTML page that contains a description of the items in a shopping catalog. 204 is a database which may store the shopping catalog information. This may be queried by the content handler in order to create the HTML page. If the request needs to be throttled then the web server hands over the request to the content throttler (202). The content throttler determines whether a request should be processed or not based on a combination of metrics such as load on the content handler, load on the database, load on the server that is running the application components, available bandwidth etc. See figure 6 for a sample scenario of content based throttling based on the load on a backend database. If the content throttler determines that the request should not be processed then it sends back a custom page (this page could contain information that describes that the web site is experiencing peak load and the user at the browser could try a while later) to the web server. If the content throttler determines that the request could be admitted then it hands the request to the content handler to process the request. The web server obtains the content that is obtained by the content handler or the content throttler and sends the response back to the browser that originated the request.

Fig 5 is a flow chart for the step that determines if a request needs to be subjected to content based throttling. The web server determines the type of the request based on the URL in the request and the web server configuration. There is a mapping in the web server configuration that contains the plugins that perform additional processing on request of certain types. The web server checks if the request type is flagged to be processed by the content based throttler. If the flag is set then it is determined that this request needs to be subjected to content based throttling.

Fig 6 is a flow chart depicting the decision process for content based throttling. The request is identified by making use of the configuration information in the web server configuration and matching it against the load parameter that is monitored (in this case the request is identified as being for a backend database). The request the request is associated with the response time as the load parameter (step 502). There is a threshold for the response time that has been set either at configuration time or dynamically by a component that is responsible for determining the load on the database. The content throttler determines whether the request should be admitted or not based on the current value of the response time (step 503). If the current response time is less than the threshold that has been set then the request is admitted, otherwise the request is dropped. Another example of a content throttler would be based on the size of the output being generated for requests to images stored in a repository. The throttler obtains the size of

the image from the repository and if the size of the image is less than a threshold size it may decide to admit the request, otherwise it may choose to send a custom HTML page refusing the admission of request. Note that these are only a few of the algorithms for throttling. Several other algorithms could be used to arrive at the throttling decision.

① → The embodiment described above is distributed in nature as the throttling is performed at every web server in the system. This is different from the current art where the admittance control decisions are made at the level of L4LB which is a central entity. Content based routing (CBR) which is sometime implemented on top of L4LB routes requests to web servers based on the content. Implementations of CBR make use of a set of rules to match URL's with the rules and tend to be bottlenecks as they inspect every request that arrives to the web site [3].

### Claims

- ② →
1. A method of controlling the admittance of requests to the processing components by the content throttler by differentiating the type of request based on the content in a distributed heterogeneous computing environment.
  2. A method of claim 1 where the requests can be prioritized within the same type based on further refinement of the content.
  3. A method of claim 1 where the dependencies of the processing components are represented.
  4. A method of claim 3 where the load on the processing components can be aggregated centrally to determine the load imposed on a dependee processing component.
  5. A method of claim 4 where the admittance control to dependent components is performed by determining the load on the dependee component.
  6. A method of claim 1 where the rejected requests could have a user defined response associated with them. (e.g. If a request is rejected a user defined page which indicates "Server Busy" could be sent)
  7. A method of claim 1 where throttling can be applied at any tier to control admittance to the next processing component along the request flow path.
  8. A method of claim 3 where the admittance of requests could be prevented to a dependent processing component if a dependee processing component is not available.
  9. A method of claim 1 where the normal response to a request could be altered by a content throttler based on the load of the processing components. (e.g. Consider the case when outbound bandwidth for a customer segment has been exceeded. A request for a very large image file which does not provide monetary value to the response could be replaced by a smaller text file which describes the image)
  10. A method of claim 1 where the throttling functionality is distributed.

### References

1. US5918017: System and method for providing dynamically alterable computer clusters for message routing
2. "Highly Scalable System and Method of Regulating Internet Traffic to Server Farm to Support (min,max) Bandwidth Usage Based Service Level Agreements", IBM Docket

IBM Confidential

No. YO9990470. Patent filed on 4/5/2000

3. IBM Websphere Performance Pack: Caching and Filtering with IBM Web Traffic Express. IBM Red Book SG24-5859-00.



IBM Confidential

☒ ABC

YOR8-2000-0394

☒ Patentability☐ Clearance

Functional Code \_\_\_\_\_

☐ Validity

PTAC No. \_\_\_\_\_

☐ State of Art**Search Request**Title of Invention: **A METHOD, SYSTEM AND PROGRAM PRODUCTS****FOR DISTRIBUTED CONTENT BASED THROTTLING IN A**Feature(s) to be Searched: **COMPUTING ENVIRONMENT**

Throttling requests to servers in a clustered hierarchy of servers based on the content of the request and performance metrics of the servers.

Altering the response to a server request to accommodate resource constraints of an end user.

See attached cbt.lwp file and invention disclosure.

Keywords: request throttling, request admittance or admission

Additional Information Including Known Prior Art: \_\_\_\_\_

Requested by:  Douglas W. Cameron

Location: 39-254, IPL Yorktown

Date: 6/13/00

Close Art:

**SEARCH REPORT**

Other Art:

Field of Search:

Other Comments:

bc: BR

From:

Date:

**EXHIBIT 2**

800-01480

4-FILE COPY-DEPARTMENT CONDUCTING/COORDINATING SEARCH

RECEIVED JUL 06 2008

Searcher: Jim Kunkle

IBM CONFIDENTIAL

- Search Request No: YO8-00-0394 (Patentability)
- Search Request For: Douglas W. Cameron

**DESCRIPTION:**

Throttling requests to servers in a clustered hierarchy of servers based on the content of the request and performance metrics of the servers.

Altering the response to a server request to accommodate resource constraints of an end user.

**FIELD OF SEARCH**

CLASS	SUBCLASSES
709	202, 203, 218, 219, 223, 226, 227, 228, 229

**DATABASES**

X	WPAT	X	JAPIO	X	APS
X	TDBS	X	DOSS		WPI
	OTHER				

**RELATED ART**

PATENT NUMBER	PATENT NUMBER	PATENT NUMBER	IBM-TDB	LITERATURE
6006264				
6067559				
5926798				

**BACKGROUND ART**

PATENT NUMBER	PATENT NUMBER	PATENT NUMBER	IBM-TDB	LITERATURE
6073241				
5864870				
5961602				

Exhibit 3

(page 1 of 2)

How does each of the Related Art references correspond to the Search Feature? (Claims if Validity/Clearance)

Patent 6006264 relates to a system for control of information flow between client and server using a "best fit" selection among multiple web servers. The references cited against this patent were checked.

Patent 5926798 discloses a server selection based on content and business policies of each server.

Patent 6067559 relates to server architecture which includes application managers which select agents from a group. The references cited against this patent were checked.

What features does the Background Art provide:

Patent 6073241 -- web browser tracking across distinct WWW domains.

Patent 5864870 -- server storage of formats for matching of requests.

Patent 5961602 -- web server selection based on web activity level.

What features were not found?

Comments:

Signature

J. Q. Kunkle

Date returned to WAIPL

7/6/00

Exhibit 3  
(page 2)

Name

Internal Zip

City & State

To: Srirama M. Krishnakumar

H4-E54

Hawthorne

*File*

Classification: IBM Confidential

Prepared by IBM Attorney

From

Date: July 12, 2000

Name & Tie/Ext.: Douglas W. Cameron 862-3244

Title/Dept. Name: G2WC - Intellectual Property Law Department

Location/Internal Zip: 39-254/T.J. Watson Research Center, Yorktown

Mail Address: P O Box 218, Yorktown Heights, NY 10598

Electronic Address: NAC at IBMUSM11

Subject: Request for Review of Art Related to Disclosure YOR8-2000-0394  
A METHOD, SYSTEM AND PROGRAM PRODUCTS FOR DISTRIBUTED CONTENT  
THROTTLING IN A COMPUTING ENVIRONMENT

Reference: USPs5,926,798, 6,067,559, 5,961,602, 5,864,870, 6,073,241, 6,006,264

Please review the above-referenced patents which were found during a patentability search directed subject disclosure. Upon completion of your review, please provide me with a memo detailing the differences between your invention and each of the referenced patents.

Note particularly US Patent 5,926,798 to Carter, which discloses an intelligent agent which receives requests from clients and makes decisions as to which server should serve a request based upon content related information and business policies.

Your prompt response would be appreciated.

  
Douglas W. Cameron

DWC/lf

Enclosures

cc: Liana L. Fong w/o Enc.  
Karen Appleby w/o Enc.  
Donald P. Pazel w/o Enc.  
German Goldszmidt w/o Enc.

H4-C08  
H4-A14  
H2-B15  
H4-B18

Hawthorne  
Hawthorne  
Hawthorne  
Hawthorne

*EXHIBIT 4*



cc: German Goldszmidt/Watson/IBM@IBMUS, Liana L Fong/Watson/IBM@IBMUS, Donald P Pazel/Watson/IBM@IBMUS, Karen Appleby-Hougham/Watson/IBM@IBMUS  
From: Douglas W Cameron/Watson/IBM@IBMUS  
Subject: Re: \*IBM Confidential: Comments on the art related to disclosure YOR8-2000-0394 A method, system and program products for distributed content based throttling in a computing environment

I am trying to get a the point of novelty in view of the art found, hence the questions below? Once we get the point of novelty, we can then see what we may claim in a patent.

Referring to your comments on US Patent 5, 926, 798, you mention that both this patent and CBT select a server which is a best fit for the content requested, but that server selection in the patent is different from CBT.

What is that difference?

What do you mean by the statements below that: "CBT is used to throttle requests to the backend based on load rather than select servers based on performance metrics. CBT is applied at the server after the server selection has been done." On its face, this statement appears to contradict the claims section of the attached file?

Referring to your comments on US Patent 6,067,559, which deals with steering requests in a hierarchy, see second bullet of item 1, and item 5 in the claims section of the attached file.

- Each server determines the admittance of requests to the server in the next tier based on the content of the request and the collected metrics.

5. Throttling can be applied at various tiers to control admittance to the servers in the next tier in the request flow path in a recursive fashion.



Srirama Krishnakumar  
07/31/2000 02:49 PM

To: Douglas W Cameron/Watson/IBM@IBMUS  
cc: German Goldszmidt/Watson/IBM@IBMUS, Liana L Fong/Watson/IBM@IBMUS, Donald P Pazel/Watson/IBM@IBMUS, Karen Appleby-Hougham/Watson/IBM@IBMUS  
From: Srirama Krishnakumar/Watson/IBM@IBMUS  
Subject: \*IBM Confidential: Comments on the art related to disclosure YOR8-2000-0394

Hi Doug,  
Following are the comments to the patents that you had sent us. Please let us know the next course of action. I would also like to let you know that we have submitted some parts of the project description to be published in a paper in December. I guess that we will have the process completed before that.  
Cheers  
Srirama

Patent number: 6,073,241

EXHIBIT 5

page 1 of 2

- This patent relates to tracking users at web browsers using Cookies. It has no relation to throttling requests at a web server.

Patent number: 5, 864, 870

This is a patent on an object oriented database that can be accessed from a network. It is like an Internet file system. Again this is not related to Content Based Throttling (CBT).

Patent number: 6,067,559

Patent on an architecture for steering requests in a heirarchy. This appears to be a patent on a web server such as IIS (considering that the patent has been filed by Microsoft). There is no reference here on CBT.

Patent number: 5,961,602

Patent on a web crawling tool that gathers data from servers on the internet during off peak times. Not related to CBT.

Patent number: 5, 926, 798

Patent on an agent used by a client to select the server which would be best fit for the content requested. Server selection is different from CBT.


Patent number 6,006,264

This might be the most relevant prior art. This patent relates to a content aware switch which is used to select the server based on the content. This functionality is similar to the Network Dispatcher which selects the most appropriate server based on the content and load. CBT is used to throttle requests to the backend based on load rather than select servers based on performance metrics. CBT is applied at the server after the server selection has been done. This is the difference between CBT and the idea mentioned in this patent.

EXHIBIT 5

Page 2 of 2

Srirama Krishnakumar  
08/01/2000 01:25 PM

To: Douglas W Cameron/Watson/IBM@IBMUS  
cc: German Goldszmidt/Watson/IBM@IBMUS, Liana L Fong/Watson/IBM@IBMUS, Donald P Pazel/Watson/IBM@IBMUS, Karen Appleby-Hougham/Watson/IBM@IBMUS  
From: Srirama Krishnakumar/Watson/IBM@IBMUS  
Subject: Re: \*IBM Confidential: Comments on the art related to disclosure YOR8-2000-0394 A method, system and program products for distributed content based throttling in a computing environment   
Importance: Normal

Hi Doug,

Let me first start with a brief description of CBT which explains it's novelty and then elaborate the difference between CBT and prior art described in U.S Patents 5, 926, 798 and 6,067,559. Both of these patents describe a process of selecting servers to handle the requests.

CBT is used on a web server or any application server to inspect the content of the request and make a decision about whether the request is to be processed or not. e.g if the request is for an image that is very large in size then CBT could be used to throttle admittance of requests for this image. The throttling may be driven policies that are based on system load, available bandwidth etc. In some cases the requests are sent to backend servers for further processing. **If the backend server is overloaded then CBT may be used to throttle requests to the backend server such that the backend server is protected from overload. This is the novelty of CBT. CBT does not select servers. It only determines whether a request should be admitted to the backend or not.**

Patent 5, 926, 798 is used to select a server which provides the best chance of serving a particular request. In this case there is an agent on the client side (note that a client is a software program such as a browser that is used by the end user who is making the request) which inspects the content of the request and tries to match the content to the servers (note that these servers may be different web sites) that may be best suited for handling the request (an example of this type of agent would be a web crawling robot that could suggest "priceline.com" or "yahoo travel" when the content of the request is for an airline ticket). Note that this is implemented on the **client side** and is useful for only selecting servers. The only similarity between this patent and CBT is that both inspect the content of the request. The use of content in both cases is very different.

Patent 6,067,559 is used to select servers based on the load on the servers. Even in this case the only similarity is that both CBT and the art described use content. But in CBT content is used to determine whether the request can be admitted or not, while in the art described content is used to determine the server that should handle the request.

Cheers  
Srirama

Douglas W Cameron  
08/01/2000 08:26 AM

To: Srirama Krishnakumar/Watson/IBM@IBMUS

EXHIBIT 6

Lucille Fourgeot  
02/09/2001 01:54 PM

To: Srirama Krishnakumar/Watson/IBM@IBMUS  
cc: Liana L Fong/Watson/IBM@IBMUS, Karen Appleby-Hougham/Watson/IBM@IBMUS, Donald P Pazel/Watson/IBM@IBMUS, German Goldszmidt/Watson/IBM@IBMUS  
From: Lucille Fourgeot/Watson/IBM@IBMUS  
Subject: Yorktown Invention Disclosure YOR8-2000-0394 - A METHOD, SYSTEM AND PROGRAM PRODUCTS FOR DISTRIBUTED CONTENT THROTTLING IN A COMPUTING ENVIRONMENT

In order that I may prepare a patent application for the subject disclosure, please answer the following questions or provide information requested below:

1. Please define what is meant by "the content of a request". The scope of this term will be used to distinguish your invention over what is known in the art.
2. Can I assume that "tiered server farms" are well known in the art?
3. Referring to the second paragraph of page 2, is the "payload of the data" the same thing as the "content of a request"?
4. Referring to the third paragraph of page 2, what does EJB stand for?
5. Referring again to the third paragraph of page 2, is "priority based throttling" known in the art?
6. If priority based throttling is known in the art, how is the priority determined?
7. Referring to the first paragraph on page 3, do the terms "JServ" or "TCat" have a full name?
8. Could you graphically illustrate a URL with the request type and the Web Server Configuration information that are both used to determine whether a requester needs to be subjected to content based throttling.
9. How is the request type J in the patent application for YO9-1999-0470 (Highly Scalable System and Method of Regulating Internet Traffic to Server Farm to Support (MIN,MAX) Bandwidth Usage-Based Service Level Agreements) related to the request of your invention?
10. In the patent application for the '470 docket, will a request of type J be rejected from the server farm if sufficient bandwidth is not available?
11. How is the flag set?
12. Doesn't the '470 docket teach admitting a request to the server farm based on the request type where the request will be denied if the target rate for the request type is exceeded?
13. Referring to the first sentence of the last paragraph of page 4, is the request identified by making use of configuration information or is it identified by information in the requested itself?

Proposed claim:

1. A method of controlling internet server farm traffic admitted to a server farm, said method comprising:  
when a request arrives at an access point of said farm, determining the type of said request;  
determining whether said request should be throttled based upon said type of request;  
if said request is to be throttled, admitting said request to said farm for further processing only if said request compares in a predetermined relationship to load parameters of said farm.
2. A method as recited in claim 1, for the comprising automatically admitting said request to said server farm if it is determined that said request need not be throttled.
3. A method as recited in claim 1, wherein said request is an hypertext transfer protocol request from a WEB Browser.

EXHIBIT 7

page 1 of 2

4. A method as recited in claim X, wherein said hypertext transfer protocol request is a request for a HTML file which is stored in a static repository.

5. A method as recited in claim 1, wherein said request is a request for a HTML file which must be dynamically configured in response to said request.

6. A method as recited in claim Y, wherein said dynamic request is of a request type that is to be throttled.

Lucille A. Fourgeot for  
Douglas W. Cameron  
Intellectual Property Law  
39-262 - Yorktown  
T/L 862-3265; (914) 945-3265  
Fax: (914) 945-3281, 945-4249  
fougeotl@us.ibm.com

Exhibit 7  
page 2 of 2

To: Douglas W Cameron/Watson/IBM@IBMUS  
cc: Liana L Fong/Watson/IBM@IBMUS, Karen Appleby-Hougham/Watson/IBM@IBMUS, Donald P Pazel/Watson/IBM@IBMUS, German Goldszmidt/Watson/IBM@IBMUS

From: Srirama Krishnakumar/Watson/IBM@IBMUS  
Subject: \*IBM Confidential: Re: Disclosure YOR8-2000-0394 - A METHOD, SYSTEM AND PROGRAM PRODUCTS FOR DISTRIBUTED CONTENT THROTTLING IN A COMPUTING ENVIRONMENT Request for Clarification

Importance: Normal

Hi Doug,

I received and read the Docket YO999-374. I would like to make the following broad observations on the similarities and the differences between content based request control (CBRC) in our invention and the Communications Bandwidth Manager (CBM).

1. CBM is concerned with the shaping of traffic at a **central** admittance point while CBRC is a **distributed** scheme.
2. CBM is applicable only to bandwidth management as it is admitting requests based on the outbound traffic. CBRC is applicable to a broader class of load management. e.g. CBRC can be used to control load on a backend database and prevent server crashes.
3. While the invention described in Docket YO999-374 refers to the terms "IP address to URL's" it does so in the context of sorting requests based on IP addresses and URL's. This sorting is done based on amount of traffic generated. CBRC sorts requests based on the application components that handle the request. CBRC does not limit itself to the URL's. Content is the complete payload of the request and includes parameters passed along with the URL such as query strings, content of HTTP POST etc.
4. The CBM scheme decides whether the request should be admitted or not. CBRC delegates this responsibility to the content throttler (component 202 in figure 3) which makes the determination. The content throttler could be different for each type of request instead of being a single central entity.

Please let me know if you have any further comments.

Cheers  
Srirama

Douglas W Cameron  
02/13/2001 08:34 AM

To: Srirama Krishnakumar/Watson/IBM@IBMUS  
cc: Liana L Fong/Watson/IBM@IBMUS, Karen Appleby-Hougham/Watson/IBM@IBMUS, Donald P Pazel/Watson/IBM@IBMUS, German Goldszmidt/Watson/IBM@IBMUS

From: Douglas W Cameron/Watson/IBM@IBMUS  
Subject: \*IBM Confidential: Disclosure YOR8-2000-0394 - A METHOD, SYSTEM AND PROGRAM PRODUCTS FOR DISTRIBUTED CONTENT THROTTLING IN A COMPUTING ENVIRONMENT Request for Clarification

Docket YO999-374 teaches using content data (URL's) to throttle data to a server farm?  
Can you clarify. I just want to make sure I have the invention down right before I start writing up the application.

See the last paragraph of page 5 of Docket YO999-374, which teaches the classifying and shaping of traffic using various levels from the Internet Protocol (IP) addresses to User Resource Locators (URLs). Also in the second paragraph of page 8 of the '374 Docket, it states that traffic is classified and shaped using various levels such as from the Internet Protocol (IP) addresses to User Resource Locators (URLs). Then, on page 7 of the same docket, it states that the incoming traffic is classified into various class/type queues, and each

EXHIBIT 8 page 1 of 2

queue has selective water levels to trigger selective dropping of newly arriving packets, where information on outbound link bandwidth usage is used.

EXHIBIT 8

Page 2 of 2

Please let me know if you have any further comments.

Cheers

Srirama

Douglas W Cameron  
02/13/2001 08:34 AM

To: Srirama Krishnakumar/Watson/IBM@IBMUS  
cc: Liana L Fong/Watson/IBM@IBMUS, Karen Appleby-Hougham/Watson/IBM@IBMUS, Donald P Pazel/Watson/IBM@IBMUS, German Goldszmidt/Watson/IBM@IBMUS  
From: Douglas W Cameron/Watson/IBM@IBMUS  
Subject: \*IBM Confidential: Disclosure YOR8-2000-0394 - A METHOD, SYSTEM AND PROGRAM PRODUCTS FOR DISTRIBUTED CONTENT THROTTLING IN A COMPUTING ENVIRONMENT Request for Clarification

Docket YO999-374 teaches using content data (URL's) to throttle data to a server farm?  
Can you clarify. I just want to make sure I have the invention down right before I start writing up the application.


See the last paragraph of page 5 of Docket YO999-374, which teaches the classifying and shaping of traffic using various levels from the Internet Protocol (IP) addresses to User Resource Locators (URLs). Also in the second paragraph of page 8 of the '374 Docket, it states that traffic is classified and shaped using various levels such as from the Internet Protocol (IP) addresses to User Resource Locators (URLs). Then, on page 7 of the same docket, it states that the incoming traffic is classified into various class/type queues, and each queue has selective water levels to trigger selective dropping of newly arriving packets, where information on outbound link bandwidth usage is used.

EXHIBIT 9



394

EXHIBIT 11


 Srirama Krishnakumar  
03/15/2001 10:54 AM

To: Douglas W Cameron/Watson/IBM@IBMUS  
cc: Liana L Fong/Watson/IBM@IBMUS, Karen Appleby-Hougham/Watson/IBM@IBMUS, Donald P Pazel/Watson/IBM@IBMUS, German Goldszmidt/Watson/IBM@IBMUS  
From: Srirama Krishnakumar/Watson/IBM@IBMUS  
Subject: \*IBM Confidential: Re: Disclosure YOR8-2000-0394 - A METHOD, SYSTEM AND PROGRAM PRODUCTS FOR DISTRIBUTED CONTENT THROTTLING IN A COMPUTING ENVIRONMENT Request for Clarification

Hi Doug,  
I was wondering if you had a chance to look at the clarifications we made.  
Cheers  
Srirama

----- Forwarded by Srirama Krishnakumar/Watson/IBM on 03/15/2001 10:52 AM -----

EXHIBIT 10

 Srirama Krishnakumar  
02/22/2001 04:06 PM

To: Douglas W Cameron/Watson/IBM@IBMUS  
cc: Liana L Fong/Watson/IBM@IBMUS, Karen Appleby-Hougham/Watson/IBM@IBMUS, Donald P Pazel/Watson/IBM@IBMUS, German Goldszmidt/Watson/IBM@IBMUS  
From: Srirama Krishnakumar/Watson/IBM@IBMUS  
Subject: \*IBM Confidential: Re: Disclosure YOR8-2000-0394 - A METHOD, SYSTEM AND PROGRAM PRODUCTS FOR DISTRIBUTED CONTENT THROTTLING IN A COMPUTING ENVIRONMENT Request for Clarification

Importance: Normal

Hi Doug,  
I received and read the Docket YO999-374. I would like to make the following broad observations on the similarities and the differences between content based request control (CBRC) in our invention and the Communications Bandwidth Manager (CBM).

1. CBM is concerned with the shaping of traffic at a **central** admittance point while CBRC is a **distributed** scheme.
2. CBM is applicable only to bandwidth management as it is admitting requests based on the outbound traffic. CBRC is applicable to a broader class of load management. e.g. CBRC can be used to control load on a backend database and prevent server crashes.
3. While the invention described in Docket YO999-374 refers to the terms "IP address to URL's" it does so in the context of sorting requests based on IP addresses and URL's. This sorting is done based on amount of traffic generated. CBRC sorts requests based on the application components that handle the request. CBRC does not limit itself to the URL's. Content is the complete payload of the request and includes parameters passed along with the URL such as query strings, content of HTTP POST etc.
4. The CBM scheme decides whether the request should be admitted or not. CBRC delegates this responsibility to the content throttler (component 202 in figure 3) which makes the determination. The content throttler could be different for each type of request instead of being a single central entity.

**Douglas W Cameron**

03/26/2001 09:02 AM

To: Srirama Krishnakumar/Watson/IBM@IBMUS  
cc: Liana L Fong/Watson/IBM@IBMUS, Karen Appleby-Hougham/Watson/IBM@IBMUS, Donald P Pazel/Watson/IBM@IBMUS, German Goldszmidt/Watson/IBM@IBMUS  
From: Douglas W Cameron/Watson/IBM@IBMUS  
Subject: \*IBM Confidential: Re: Re: Disclosure YOR8-2000-0394 - A METHOD, SYSTEM AND PROGRAM PRODUCTS FOR DISTRIBUTED CONTENT THROTTLING IN A COMPUTING ENVIRONMENT Request for Clarification

Importance: Normal

Sorry about the extensive delay. In your note below of 2/22 in which you compare your invention with Docket YO999-374, you mention that your invention is a distributed scheme and that Docket YO999-374 sorts requests based on IP addresses and URL's. [REDACTED]

Give me a call so we can formulate the exact picture of your invention and file a patent application, if we have something new to file on. Then, I'll either file a patent application myself or send it to outside counsel.

Douglas W. Cameron  
Attorney, Intellectual Property Law  
IBM Research  
Internet: nac@us.ibm.com  
Voice: (TL 862) 914-945-3244  
Admin (Lucille Fourgeot): x3265  
Fax: x3281

PREPARED BY IBM ATTORNEY / PRIVILEGED AND CONFIDENTIAL

Srirama Krishnakumar  
03/15/2001 10:54 AM

To: Douglas W Cameron/Watson/IBM@IBMUS  
cc: Liana L Fong/Watson/IBM@IBMUS, Karen Appleby-Hougham/Watson/IBM@IBMUS, Donald P Pazel/Watson/IBM@IBMUS, German Goldszmidt/Watson/IBM@IBMUS  
From: Srirama Krishnakumar/Watson/IBM@IBMUS  
Subject: \*IBM Confidential: Re: Disclosure YOR8-2000-0394 - A METHOD, SYSTEM AND PROGRAM PRODUCTS FOR DISTRIBUTED CONTENT THROTTLING IN A COMPUTING ENVIRONMENT Request for Clarification

Hi Doug,  
I was wondering if you had a chance to look at the clarifications we made.  
Cheers  
Srirama

----- Forwarded by Srirama Krishnakumar/Watson/IBM on 03/15/2001 10:52 AM -----

 Srirama Krishnakumar  
02/22/2001 04:06 PM



International Business Machines Corporation

Thomas J. Watson Research Center  
P.O. Box 218  
Yorktown Heights, NY 10598  
914/945-1255

RECEIVED

April 4, 2001

APR 6 - 2001

Sean M. McGinn Esq.  
McGinn & Gibb, PLLC  
8321 Old Courthouse Road, Suite 200  
Vienna, Virginia 22182-3817

McGINN & GIBB, P.C.

Re: Preparation and Filing of Patent Application for YOR920010330US1  
Yorktown Disclosure No. YOR8-2000-0394

Title: METHOD, SYSTEM AND PROGRAM PRODUCTS FOR DISTRIBUTED  
CONTENT THROTTLING IN A COMPUTING ENVIRONMENT

Dear Sean:

The materials required for preparation of a patent application for the above-referenced docket including an original invention disclosure, a copy of an embodiment and drawings, copies of two patent applications previously filed, YO999-374 and YO999-470, a search report with references, correspondence with inventors, and a diskette with a copy of the drafted embodiment. Please prepare the application in accordance with the IBM Outside Counsel Instructions, as you have done for us previously. Please let me know, before starting work on these application, an approximate time and cost estimate, and please add a Beauregard claim if applicable.

Please conduct the work directly with the inventors while keeping our office informed about scope of coverage and financial decisions.

When the application is in final form, please prepare all formal papers and file the application in the USPTO by May 30, 2001. The formal papers are to be prepared by your office, listing the names of all Yorktown attorneys on the Declaration and Power of Attorney as follows:

Manny W. Schecter (Reg. 31,722), Lauren C. Bruzzone (Reg. 35,082), Christopher A. Hughes (Reg. 26,914), John E. Hoel (Reg. 26,279), Joseph C. Redmond, Jr. (Reg. 18,753), Marc A. Ehrlich, (Reg. 39,966), Derek S. Jennings (Reg. 41,473) Stephen C. Kaufman (Reg. 29,551), Richard M. Ludwin (Reg. 33,010), Robert P. Tassinari (Reg. 36,030), Robert M. Trepp (Reg. 25,933), Louis P. Herzberg (Reg. 41,500), Daniel P. Morris (Reg. 32,053), Paul J. Otterstedt (Reg. 37,411) Marian Underweiser (Reg. 46,134), Douglas W. Cameron (Reg. 31,596) and Gail H. Zarick (Reg. 43,303).

Send correspondence to: Outside Counsel  
Direct Telephone Calls to: Outside Counsel

EXHIBIT 13  
page 1 of 2

Page 2  
April 3, 2001

Also, your formal papers should be sent directly to the inventors for execution with no notarization being required. The inventors are responsible for mailing the formal papers directly back to outside counsel.

The "Taiwan Oath & Assignment" form must be completed. The form must have all the required information completely filled in and must be signed and dated by all named inventors in the subject patent application. When the form is complete, and all inventor(s) signatures and dates have been obtained, please forward the original Taiwan Oath & Assignment back to my office.

Please forward a copy of all papers (application, drawings, formal papers, etc.) to this office, with a diskette (5.0) in Lotus WordPro.

The primary inventor for you to contact is Srirama M. Krishnakumar who may be reached at (914) 784-7945 at our Hawthorne location.

The application may be filed with informal drawings; however, formal drawings will be required shortly thereafter in order to permit timely distribution of the applications to our World Trade locations. Please have your office prepare the formal drawings and forward to us one set no later than one (1) month after the filing date. You would handle filing of the formal drawings in the U.S. Patent Office. (please note that our fee addressee should be our office in Yorktown Heights.)

If possible, please send a single invoice only after the applications have been filed. The invoice should cover all your fees, together with application USPTO filing fees and other disbursements to cover preparation of the applications and drawings; as well as, minor other items relating to the preparation and filing of the applications.

Please send detailed billing statements with invoices. The statements should include date(s) when services were rendered, service(s) rendered, time spent, by whom, and at what billing rate. Payment of invoices may be delayed when a billing statement is not received.

If you have any questions or comments, please feel free to contact me.

Best regards,

Douglas W. Cameron  
Intellectual Property  
Law Department  
(914) 945-3244

DWC/lf  
Enclosures

c: S. M. Krishnakumar, L.L. Fong, K. C. Appleby-Hougham, D. P. Pazel, G. Goldszmidt,  
L. C. Bruzzone and B. Rasa

EXHIBIT 13

Page 2 of 2

LAW OFFICES OF  
**McGINN & GIBB, PLLC**

A PROFESSIONAL LIMITED LIABILITY COMPANY  
PATENTS, TRADEMARKS, COPYRIGHTS, AND INTELLECTUAL PROPERTY LAW  
8321 OLD COURTHOUSE ROAD, SUITE 200  
VIENNA, VIRGINIA 22182-3817  
TELEPHONE: (703) 761-4100  
FACSIMILE/DATA: (703) 761-2375; 761-2376  
E-MAIL: MCGINNGIBB @ AOL.COM

SEAN M. MCGINN  
FREDERICK W. GIBB, III  
PHILLIP E. MILLERT  
FREDERICK E. COOPERRIDER†  
CHRISTOPHER N. SEARST  
MATTHEW DYOR  
CHRISTOPHER M. MCGINN\*

†MEMBER OF BAR OTHER THAN VA  
\*PATENT ENGINEER (NON-ATTORNEY)

April 10, 2001

**FACSIMILE TRANSMISSION**

To: Douglas W. Cameron, Esq.  
Attorney, I. P. Law Dept.  
Thomas J. Watson Research Center

FAX No.: (914) 945-3281

From: Sean M. McGinn

FAX No.: (703) 761-2375

Re: IBM Docket No. YOR920010330US1  
**Filing Due Date: May 30, 2001**

No. of Pages (including this cover page): 2

Dear Douglas:

Thank you for your letter dated April 4, 2001, and the enclosed disclosure materials, in the above docket, for which we thank you very much.

We should have no problem filing the final draft patent application on our end, by May 30, 2001. We understand that there are no bar dates presently in the disclosure.

I estimate that we can prepare the patent application for approximately \$4100-\$4800, and the charges for preparing the formal papers, obtaining execution and filing the case to be about \$390 additional to the application preparation costs mentioned above. Besides the above, this fee would also include reviewing and reporting the Official Filing Receipt, reviewing and reporting the Assignment Recordation Document, preparing and filing an IDS (if relevant art is identified), and Submission of Formal Drawings (but not the cost of the drawings themselves). Please let me know if this estimate is acceptable.

We will prepare the formal drawings and forward them to you shortly after the filing date.

EXHIBIT 14

Deal 1 of 2

Douglas W. Cameron, Esq.

April 10, 2001

Page 2

Shortly, we will send a first draft of the patent application to the inventors. Once again, thank you for considering our law firm.

With best regards,

Very truly yours,



Sean M. McGinn

SMM/yr

EXHIBIT 14

Page 2 of 2

Srirama,

Thanks for taking time for the telephone conversation this morning. Like I said, you guys did a really good job on the disclosure so this application should move very rapidly to filing.

As I mentioned this morning, you will probably want to start accumulating the data for the inventors so that we can be preparing the formal documents that the inventors will sign in order to execute the application. The information needed is:

- Inventor name, as desired to appear on a published patent, including, if desired any middle name(s) or initial(s)
- Inventor residence address
- Inventor mailing address, if different from residence address
- County of residence here in US
- Country of citizenship

Please review the attached first attempt at the Application. I have left some holes to be filled in as I know that the basic description and claims are on the right track. You will note that I have inserted in bold some questions and comments. You may want to respond to these questions/comments in your edited version, or you might prefer to call me. Either way is okay, but usually the inventors simply write a response in the edited version and return to me via email.

Fred Cooperrider  
McGinn & Gibb, PLLC  
ph (703) 761-2377  
email fcooperrider @ AOL.com

EXHIBIT 15

Subj: **Re: Patent Application for Distributed Content Based**  
Date: 5/25/2001 2:17:59 PM Eastern Daylight Time  
From: srirama@us.ibm.com (Srirama Krishnakumar)  
To: FCooperrider@aol.com

File: YOR300\_5\_23\_01 (38112 bytes)  
DL Time (115200 bps): < 1 minute

Hi Fred,  
Please find attached my comments. I guess we could give it a second pass if you have any further doubts. How about Tuesday (May 29) between 2PM-3PM EST?  
Cheers  
Srirama  
(See attached file: YOR300\_5\_23\_01)

FCooperrider@aol.com on 05/24/2001 06:08:48 PM

To: Srirama Krishnakumar/Watson/IBM@IBMUS  
cc:  
Subject: Patent Application for Distributed Content Based Request  
Throttling

Srirama,

Thanks for taking time for the telephone conversation this morning.  
Like

I said, you guys did a really good job on the disclosure so this application should move very rapidly to filing.

As I mentioned this morning, you will probably want to start accumulating the data for the inventors so that we can be preparing the formal documents that the inventors will sign in order to execute the application. The information needed is:

- Inventor name, as desired to appear on a published patent, including, if desired any middle name(s) or initial(s)
- Inventor residence address
- Inventor mailing address, if different from residence address
- County of residence here in US
- Country of citizenship

Please review the attached first attempt at the Application. I have left some holes to be filled in as I know that the basic description and claims are on the right track. You will note that I have inserted in bold some questions and comments. You may want to respond to these questions/comments in your edited version, or you might prefer to call me. Either way is okay, but usually the inventors simply write a response in the edited version and return to me via email.

Fred Cooperrider  
McGinn & Gibb, PLLC

EXHIBIT 16



Srirama,

Thanks for your wakeup call last night. As promised in that telecon, attached is the next version of the application on throttling. I think this version is fairly close to a final version. Note that I have made changes from the last version in bold print for your convenience, but you will probably want to read the entire text.

Also note that the claim coverage is quite expanded from the previous version. Please confirm that these claims cover all significant technical aspects of the throttling invention.

Let me know if you wish to see the figures. They are basically the diagrams you supplied.

Since we are likely quite close to being able to file this application, please expedite the transmittal to me of the inventor information so we can begin preparing the formal filing documents. This information is: inventor name (as desired to be on a printed patent; residence address; mailing address, if different from residence address; county of residence; and country of citizenship.

Fred Cooperrider  
McGinn&Gibb, PLLC  
ph (703) 761 2377  
email: fcooperrider@AOL.com

to srirama@us.ibm.com

EXHIBIT 17

Srirama,

\*Attached is the "final" version of the patent application for the web farm throttling. Please check it over and make whatever changes/corrections are required.

Also attached are the assignment, declaration, and Taiwan Oath that each inventor is to sign. Please download these three documents and have all inventors sign. They need not all sign the same copy; each can download and sign individually, if more convenient. Please return to me via Overnight Mail the executed documents. Also, please fax a copy of the executed documents to me at (703) 761-2375 or -2376 so that I have a copy in case the originals are lost in the mail.

I am also faxing a copy of the figures that will accompany the patent disclosure so that you can check everything prior to signing the documents.

Please canvas the inventors one final time to ensure that they understand that inventors and anyone else involved in the invention or preparation of the disclosure is obligated to provide any pertinent documentation to the Examiner. Have them send me a copy of anything they consider particularly relevant, such as articles, etc. that they may have in their files.

As soon as we get the executed documents back we will proceed to file this application.

Fred Cooperrider  
McGinn&Gibb, PLLC  
email: fcooperrider@AOL.com  
ph: (703) 761-2377

EXHIBIT 18

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**